Design of a duplicated fault-detecting AES chip and yet using clock set-up time violations to extract 13 out of 16 bytes of the secret key.

M. Agoyan[1,2], S. Bouquet[1,2], M. Doulcier-Verdier[1,2], J.-M. Dutertre[2], J. Fournier[1,2], J.-B. Rigaud[2], B. Robisson[1,2], A.Tria[1,2], [1]CEA-LETI-MINATEC, [2]Ecole Nationale Supérieure des Mines de Saint-Etienne, Gardanne, France.

1 Introduction

The secret keys manipulated by cryptographic circuits can be extracted using fault injections associated with differential cryptanalysis techniques [1]. Such faults can be induced by different means such as lasers, voltage glitches, electromagnetic perturbations or clock skews. Several counter-measures have been proposed such as random delay insertions, circuit duplications or error correcting codes. In this paper, we focus on an AES chip in which the circuit duplication principle has been implemented to detect fault injection. We show that faults based on clock set-up time violations can nevertheless be used to defeat the implemented counter-measure.

2 Circuit description

The Tamper-Resistant (TR) AES chip is an ASIC implementation of a secure version of the 128-bit-key AES [2] algorithm as described in [7]. The chip is implemented in HCMOS9gp 130nm STM technology and works at 50 MHz. The algorithm is an iterative one consisting of *rounds* where the input data and intermediate values are represented as 4x4 matrices of bytes. Each *round*, which is executed in 1 clock cycle, uses a different *round-key* which is iteratively derived from the input *secret key*. The security strategy used is to detect induced errors and to spread them so that the resulted erroneous *cipher-text* can no longer be used for differential cryptanalysis [3].

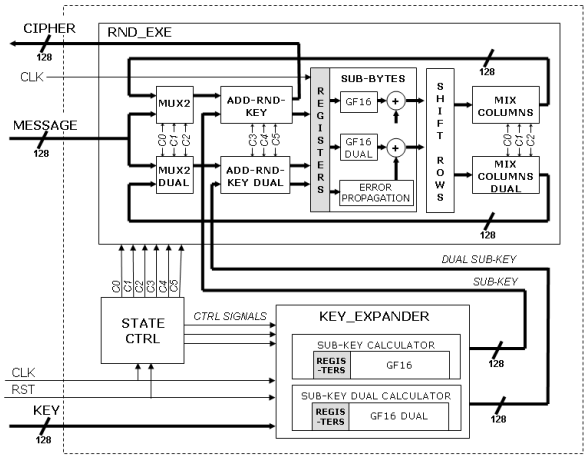

Fig. 1: TR- AES block diagram

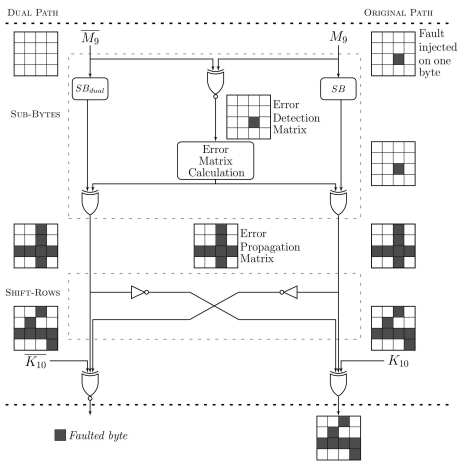

Fig. 2: Fault generated on one path with detection and propagation

Error detection is achieved by duplicating the data paths of the RND_EXE and KEY_EXPANDER blocks in a dual manner (as shown in Fig. 1), thus creating an 'original data path' working on the input message and key and a 'dual data path' working on complemented values of the input message and key. The error detection

is implemented in the SUB-BYTES module (the upper part of Fig. 2) whereby an Error Detection Matrix (EDM) is derived by comparing both data paths. Then, an Error Propagation Matrix (EPM) is built as shown in Fig. 3 before being injected on the data paths to obtain a spreading of the errors (the grey bytes in Fig. 2).

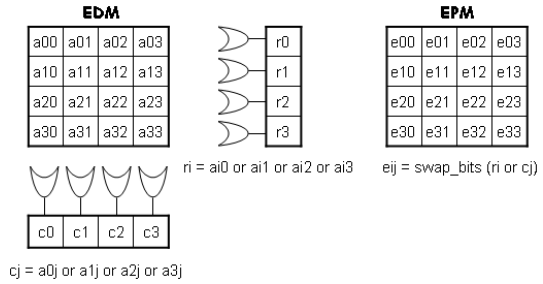

Fig. 3: Error Matrix calculation

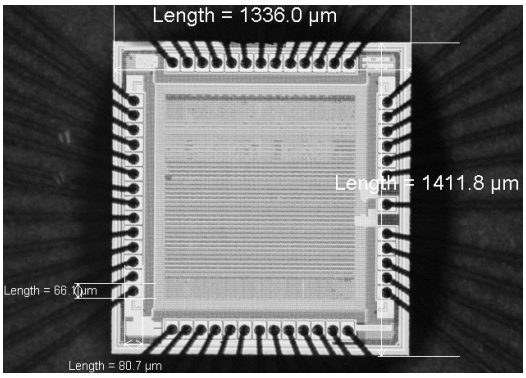

Fig. 4: Picture of the TR-AES chip

Further diffusion of the errors is achieved by cross-changing wires between the original and the dual data paths in the SHIFT_ROWS module. Fig. 2 shows what happens when an error is generated on one of the data paths at the start of the last round: the resulting *cipher-text* has too many faulty bytes to be exploited by any differential cryptanalysis technique. The resistance of this TR-AES chip against side-channel (due to the dual representations) and local laser fault (due to the error propagation mechanism) attacks is demonstrated in [7].

3 The attack scenario

Since the TR-AES chip works with an external clock, we chose to stress-test the circuit using the clock set-up time violation fault injection technique described in [4]: we decrease the clock period at a targeted cycle in order to corrupt the execution of one particular round of the AES. This faulty clock is generated using the embedded Delay Locked Loop (DLL) of a Xilinx Virtex 5 FPGA (Fig. 5).

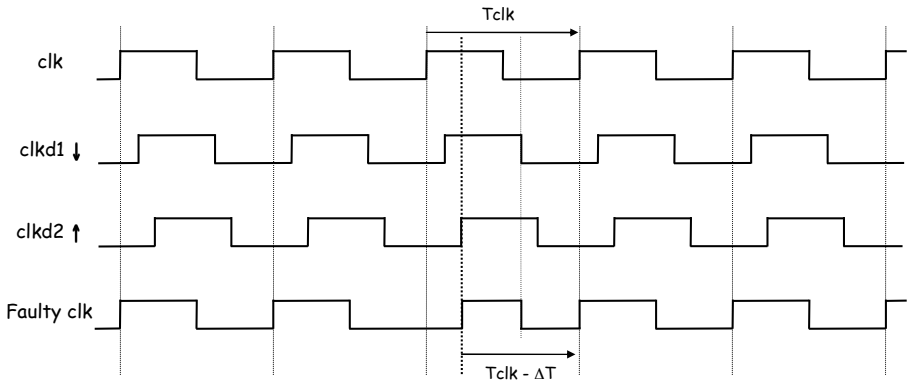

Fig. 5: Faulty Clock Signal Generation

Two clocks (*clkd1* and *clkd2*) with programmable skews are generated from the original 50MHz clock (*clk*) and the resulting faulty clock is a combination of *clkd2*'s rising edge and *clkd1*'s falling edge. The generated clock's skew ΔT varies from 0 to

ΔT_{max}=9804ps by steps of δT=76ps (the DLL's smallest elementary delay). Hence, the period of the faulty clock cycle varies between 20ns and 10.196ns. In [4], the authors generated single bit faults on a non-secure FPGA implementation of the AES by monitoring the duration of the skew.

In order to exploit the generation of single bit faults on an AES, we used two differential fault cryptanalysis techniques where bytes of the 10th *round-key* can be retrieved. In the first one, described by Giraud [5], the attacker has to inject only one bit error at the end of the 9th (one before last) round of the AES. Then by using the expected *cipher-text* and the corrupted one, bits of the secret keys can be found. For a given byte, with 3 corrupted *cipher-texts*, the corresponding sub-key byte is found with a probability of 99%. The second technique is given by Piret & Quisquater in [6]. There, a single bit fault has to be injected at the end of the 8th round of the AES such that the correct *cipher-text* and the faulty one differ by four bytes (one column). With two faulty *cipher-texts* corrupted on the same column, there are 97% chances of retrieving the corresponding column's secret *round-key*.

For those two techniques to work on the TR-AES, we need to 'bypass' the error detection. To achieve this, the same error is generated on the 'original' and the 'dual' data paths as shown in Fig. 6. Consequently the error matrix is null, the error is not detected and at the end of the *round*, only one byte is corrupted. A fault is hence generated and yet not detected by the counter-measure.

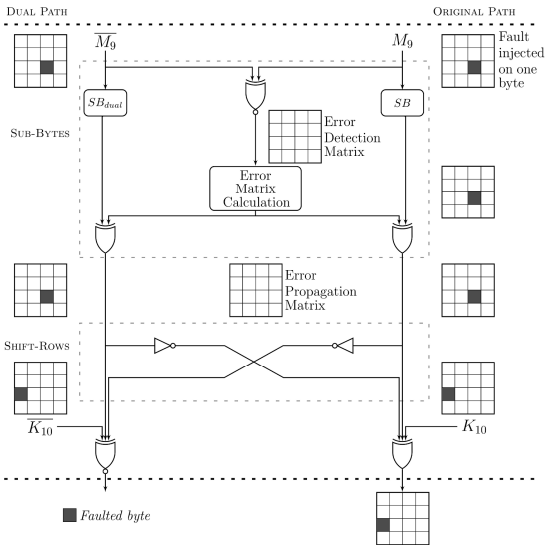

Fig. 6: Faults generated on both paths without detection nor propagation

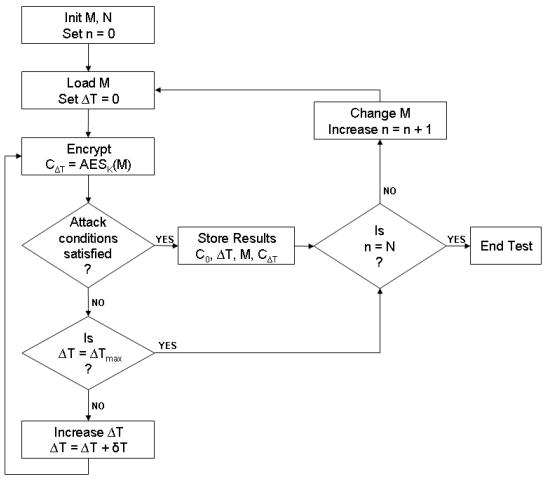

Fig. 7: Attack scheme for Giraud's or Piret-Quisquater's attacks

We used the DLL-based clock set-up time violation board to inject the desired faults. The attack scenario used is given in Fig. 7: for each of the different N random messages M, the skew on a 'specific' clock cycle (corresponding to a round) is increased until the 'attack condition' is satisfied. In the case of Giraud's attack, the clock skew is inserted during the 9th round and the 'attack condition' corresponds to having a faulty *cipher-text* with only one corrupted byte. Out of N=60000 executions,

done in 36 hours, 235 *cipher-texts* had one faulty byte and among these 6 different bytes locations were impacted. This approach revealed bytes 5, 7, 8, 9 and 10 of the AES 10^{th} *round-key* and reduced to 3 the number of possibilities for the 1^{st} one. For the Piret-Quisquater's attack, the skew is inserted during the 8^{th} round and the 'attack condition' corresponds to having an entire column of the *cipher-text* impacted. We played 20000 scenarios (in 13 hours), 9 of which induced four-byte errors on a column but only six were exploitable for the attack. We hence found bytes 2, 3, 4, 5, 6, 7, 9, 10, 12, 13, 15 and 16 of the 10^{th} *round-key*. By combining the results from both attacks we got 13 bytes of the 10^{th} *round-key* and had only 3 possibilities for a 14^{th} one. The remaining *round-key* bits could then be found using $3\times2^8\times2^8$ brute force searches. With the 10^{th} *round-key*, the original secret key can be calculated by executing the iterative *round-key* derivation algorithm in the reverse order.

4 Conclusion

The TR-AES chip is a secure cryptographic circuit having a complemented duplicated implementation designed to resist to fault attacks and to reduce the effect of side channel information leakage [7]. Using clock set-up time violations, we showed that such a counter-measure is not enough to protect against fault injections. With Giraud's and Piret-Quisquater's techniques, 13 out of 16 secret key bytes were retrieved in less than 40 hours. It is, to our best knowledge, the first reported practical results on an AES ASIC chip showing that fault detection based on duplication can be defeated using low cost techniques. This counter-measure was designed to resist to a laser beam fault injection [7], which has a local effect, as opposed to a stress induced by a clock glitch which has a global effect on both data paths. In the light of the data retrieved from the fault attacks described in this paper, we believe that combined side-channel and fault attacks can be carried on such a design.

Acknowledgements
This work was funded by the SECRICOM project (EC FP7-SEC-2007 grant 218123).

References:

[1] E. Biham & A. Shamir, "Differential fault analysis of secret key cryptosystems", in Proceedings of the Crypto'97, Lecture Notes in Computer Science, 1997.
[2] NIST, "Announcing the Advanced Encryption Standard (AES)", Federal Information Processing Standards Publication, n°19 7, November 26, 2001.
[3] M. Joye, P. Manet & J-B. Rigaud, "Strengthening hardware AES implementations against fault attack", IET Information Security, vol. 1, issue 3, September 2007.
[4] M. Agoyan, J-M. Dutertre, D. Naccache, B. Robisson, A. Tria, "When clocks fail: on critical paths and clock faults", in Proceedings of CARDIS 2010, LNCS, 2010.
[5] C. Giraud, "DFA on AES", In Advanced Encryption Standard – AES, volume 3373 of Lecture Notes in Computer Science, pages 27–41, Springer, 2005.
[6] G. Piret & J-J. Quisquater, "A differential fault attack technique against SPN structures, with application to the AES and Khazad", in Proceedings of Cryptographic Hardware and Embedded Systems, Lecture Notes in Computer Science, 2003.
[7] M. Doulcier-Verdier, J-M. Dutertre, J. Fournier, J-B. Rigaud, B. Robisson & A. Tria, *"*A side-channel and fault-attack resistant AES circuit working on duplicated complemented values*"*, in Digest of technical papers of ISSCC 2011, February 2011.